7 April 1980

NOTE FOR: Harv
Hutch

FROM: Willi

STATINTL SUBJECT: [redacted] "Methodology" for Interpreting DCID 1/2 Priorities

STATINTL 1. When Hutch first raised reservations back in November about the way [redacted] had computed is ordinal measures of the DCID's priorities, I looked at their "Appendix B" quickly, and was left with the innocuous impression that they had gone to a lot of trouble to belabor an obvious point (i.e., that not all subjects and geographical areas were proportionately represented in the different priority classes). At that time, I was more concerned over their interpretation of the rankings (i.e., the "general pattern" business) as presented in their draft Executive Summary, than with the method by which the rankings had been determined.

2. At Hutch's continued urging, I finally took a closer look at Appendix B. To put it bluntly, my reaction fell somewhere between sorrow and chagrin. What they've done is probably most charitably described as "artistic". Certainly it has no foundation in any scientific analytical principles I've ever heard of, notwithstanding the fact that they've cloaked their psuedo-calculations in high-fallutin, statistical-sounding terms like "maximum positive deviation from the norm."

3. I would surmise that the basic problem is the same one we encountered in the Executive Summary, with its vague but reassuring references to "patterns of consistency" that I, for one, couldn't begin to discern. When someone stares long enough and hard enough at a bunch of numbers, they begin to see things, things which they then try to explain in terms that sound as much as they can make them like the methods they should have used in the first place, but didn't. Rather than attempt a serious critique of what [redacted] done in Appendix B (Its hard to know where to start. And besides, Hutch has already gone that route to no avail.), I'll try the reductio ad absurdum approach. STATINTL

4. Figure I, attached, shows a hypothetical priorities matrix featuring four subjects (A, B, C and D), and the same number of priority classes (1 through 7 -- I've conveniently disregarded Hutch's valid point about the differences between priority 8's and not applicables) and class assignment totals as the actual real-life DCID. But, I've constructed my matrix in such a way that it is obviously totally and overwhelmingly dominated by Subject A, which claims 8744 of the possible 9520 assignments (or better than 9 out of every ten) including every blessed one of the Priority One's. In stark contrast, Subject D has but a single assignment.

How, you ask, can one lone measly citation (and a mediocre "Priority Three" at that) possibly be interpreted as meaning Subject D has greater relative importance than Subject A, with its fifteen Priority One's, 101 Priority Two's, 351 Priority Threes, and so forth? Easy! Just scrupulously apply the [redacted]"methodology". Via this analytical pipedream, Subject D clearly emerges as the most important in the matrix. Subject A ends up dead last in relative importance, and by quite a wide margin. STATINTL

STATINTL Absurd? Of course it is. [redacted] method has led to an STATINTL indisputably clear-cut situation being interpreted in exactly the wrong way. (Subjects B and C are intermediate cases, constructed in such a fashion as to preserve the "precise wrongness" of the results the [redacted] method produces.)

5. Where, you may next ask, did [redacted] reasoning go astray? The answer, I fear, is way back at the very first opportunity for it to do so. After correctly observing that DCID 1/2 priority intervals aren't scaled in the document itself, they conclude: "Thus, calculation of an absolute priority value assigned to a subject, or region, is not possible." Of course, its never possible to calculate "absolute priorities" (the term itself is nonsense -- all priorities are relative by definition), but that isn't what they really mean. What they're trying to say is that, since the DCID doesn't tell you precisely how much importance you should attribute to a Priority One compared to a Priority Two, or to a Two over a Three, a Three over a Four, etc., there's no basis at all for using a weighted index to evaluate relative subject emphases. Hence, a metaphysical approach like theirs is inescapable. STATINTL

"REDUCTIO AD ABSURDUM" MATRIX FOR [redacted] RANKING METHODOLOGY

FIGURE H

STATINTL

(HYPOTHETICAL)

NUMBER OF DCID 1/2 ASSIGNMENTS

SUBJECT	PRIORITY 1	2	3	4	5	6	7	TOTAL	% OF TOT.
A	15	101	351	1026	1656	2072	3523	8744	91.8
B	0	5	50	100	140	140	140	575	6.0
C	0	5	25	50	40	40	40	200	2.1
D	0	0	1	0	0	0	0	1	0.01
TOTAL	15	111	427	1176	1836	2252	3703	9520	100.0

"DEVIATIONS FROM NORM"*

SUBJECT	PRIORITY 1	2	3	4	5	6	7	TOTAL
A	+1	-1	-41	-54	-30	+4	+121	0
B	-	-2	+24	+29	+30	+4	-85	0
C	-	+3	+16	+25	+1	-7	-37	0
D	-	-	+	-	-	-	-	0

PRIORITY OF MAX. POS. DEV.	MEDIAN POSITIVE DEVIATION	RANK ORDER
7	6	4
5	4.5	3
4	3	2
3	3	1

* "DEVIATIONS FROM NORM" FOR CELL i,j EQUALS "NUMBER OF ASSIGNMENTS" MINUS THE QUANTITY "% OF TOTAL" FOR ROW i TIMES "TOTAL ASSIGNMENTS" FOR COLUMN j.

STATINTL

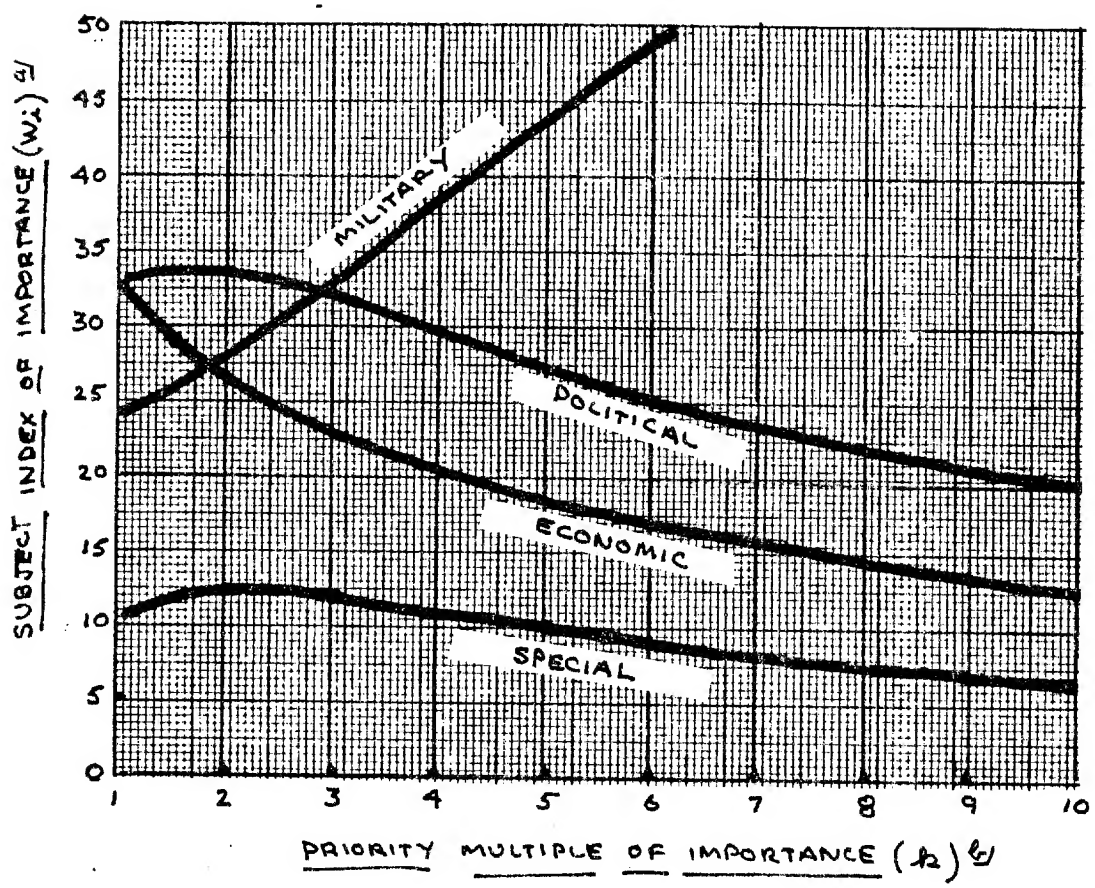
6. Fiddlesticks! In seeking to avoid the "weights" issue, [redacted] has in fact implicitly assigned equal weights to each priority level. Although we don't know for sure how much more important a Priority One should be considered compared to a Priority Two, we do know that it is more important. Otherwise, there won't be any purpose to DCID priority classes at all. Any system of weights must, to a certain extent, be arbitrary (and would be no less so if assigned by the DCID authors themselves). But that doesn't invalidate the soundness or usefulness of weighted indices for interpreting data. Moreover, in the case of the Subject Matrix at least (I haven't taken the time to examine the Region Matrices), it doesn't really matter all that much what the weighting formula is. Figure II (also attached) shows what I mean, by indicating the sensitivity of the different subjects' scores on a weighted index of relative importance to the steepness of the weighting scheme that is applied. Basically, it boils down to two cases. If you think there's a big (factor of three, or more) drop in importance in moving from one priority class to the next lower class, then the subjects rank: (1) Military, (2) Political, (3) Economic and (4) Special.

As the drop from level to level increases beyond a factor of three, Military Subjects move farther and farther away from the pack. The ordinal position of the other three Subjects within the pack does not, however, change, no matter how big the Priority differential becomes. If, on the other hand, one believes there isn't much difference (i.e., less than a factor of three) between a Priority One and a Priority Two, a Two and a Three, etc., the ordinal rankings are Military, Political and Economic equal in importance (slight differences, but too close to call in a practical sense), with Special again running a poor fourth.

STATINTL Either way, its pretty clear [redacted] is way off when it asserts that "Military and Special Subjects command the highest level of interest, with Political next and Economics well behind." This is, by any objective standard, plainly an incorrect conclusion. I haven't searched the study carefully to see how many equally false statements may have resulted from their inability to come to grips with the fact that the priorities do say something about relative importance that cannot be ignored simply because a "ratio" of importance is not provided.

SECRET

SENSITIVITY OF SUBJECT IMPORTANCE
WITHIN DCID 1/2 TO WEIGHTS ASSIGNED PRIORITY LEVELS



---

a/ $$W_i = \frac{\sum_{j=1}^{7} n_{ij} w_j}{\sum_{i=1}^{4} \sum_{j=1}^{7} n_{ij} w_j}$$

WHERE:

i DESIGNATES SUBJECT

j DESIGNATES PRIORITY

n_{ij} IS NUMBER OF MENTIONS

W_j IS PRIORITY WEIGHT

b/ $$k_2 = \frac{W_j}{W_{j+1}}$$